

ORIGINAL RESEARCH

Open Access

Estimating WUI exposure probability to a nearby wildfire



Yu Wei^{1*}, Benjamin Gannon², Jesse Young³, Erin Belval⁴, Matthew Thompson⁴, Christopher O'Connor³ and David Calkin³

Abstract

Background Estimating the factors affecting the probability of a wildfire reaching the wildland urban interface (WUI) can help managers make decisions to prevent WUI property loss. This study compiles data on fire progression, wind, landscape characteristics, and fireline built to estimate the probability of an active fire reaching nearby WUI blocks. We started by constructing funnel-shaped analysis zones between recorded fire perimeters and WUI blocks. We used zonal analysis to characterize landscape and fireline arrangement and then used a random forest modeling approach to quantify the probability of fire reaching the WUI blocks.

Results We found the probability of WUI exposure from an active fire had close relationships with several explanatory variables including wind gust velocity, suppression difficulty, control potential, fireline arrangement, road densities, WUI block sizes, and the distance between WUI and the fire's front. We found that the most important predictor variables influencing WUI exposure probability were gust, fireline arrangement, and distance from a fire ignition location to a WUI. We found that random forest models can achieve reasonable accuracy in estimating WUI fire exposure probabilities.

Conclusions Focal analyses and random forest models can be used to estimate WUI fire exposure probabilities in support of large fire suppression decisions at division to incident scales.

Keywords Fireline, Suppression, Wind gust, Random forest, Spatial analysis

Yu Wei

yu.wei@colostate.edu

¹ Colorado State University, Fort Collins, CO, USA

⁴ USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO,



^{*}Correspondence:

 $^{^2}$ USDA Forest Service, National Office, Fire and Aviation Management, Fort Collins, CO, USA

 $^{^{\}rm 3}$ USDA Forest Service, Rocky Mountain Research Station, Missoula, MT, USA

Wei et al. Fire Ecology (2023) 19:30 Page 2 of 18

Resumen

Antecedentes La estimación de los factores que afectan la probabilidad de que un incendio llegue a la interfase urbano rural (WUI en inglés) puede ayudar a los gestores de recursos a tomar decisiones para prevenir pérdidas en propiedades de la WUI. Este estudio compila datos sobre la progresión del fuego, vientos, características del paisaje, y la construcción de líneas de fuego para estimar la probabilidad de que un fuego activo llegue a las cercanías de manzanas con construcciones de esa WUI. Comenzamos por analizar zonas en forma de embudos entre registros de perímetros de fuego y manzanas de WUIs. Usamos análisis zonales para caracterizar el arreglo entre paisajes y líneas de fuego, y luego la aproximación de un modelo basado en el análisis forestal al azar para cuantificar la probabilidad de que un fuego alcance las manzanas del WUI.

Resultados Encontramos que la probabilidad de exposición de la WUI a un fuego active tuvo una relación muy estrecha con varias variables exploratorias incluyendo la velocidad de ráfagas de viento, la dificultad en la supresión, el control potencial, el arreglo espacial de la línea de fuego, la densidad de la red vial, el tamaño de las manzanas de la WUI, y la distancia entre la WUI y el frente de fuego. Encontramos que los predictores de las variables más importantes que influencian la probabilidad de exposición de la WUI fueron las ráfagas de viento, el arreglo espacial de la línea de fuego, y la distancia entre la WUI y la ubicación del punto de ignición. Encontramos que los modelos de análisis forestal al azar pueden alcanzar una exactitud razonable para estimar las probabilidades de exposición de las WUIs al fuego.

Conclusiones Los análisis focales y los modelos forestales al azar pueden ser usados para estimar las probabilidades de la exposición de las WUIs al fuego en apoyo a grandes decisiones de supresión en escalas desde división hasta de incidentes.

Background

Community protection remains a top priority for fire managers in regions with growing societal impacts from wildfire. In the western US, increasing wildfire activity and expansion of the wildland urban interface (WUI) are driving an upward trend in community exposure to wildfire and home loss (e.g., Platt et al. 2011, USDOI & USDA, 2014, Abatzoglou & Williams, 2016, Iglesias et al., 2022, Caggiano et al., 2020). For incident management teams managing a large wildfire burning toward vulnerable assets, essential information for risk-informed decision-making includes the likelihood of the fire reaching the assets and the ability for suppression to lessen that likelihood (Stratton, 2020, Calkin et al., 2021). Various fire spread models are used to estimate potential fire growth and asset exposure (e.g., Finney et al., 2011a), but limited knowledge of suppression effectiveness at division (geographical areas of firefighting operations) to incident scales has been developed (Plucinski, 2019), which raises the need for models that integrate both fire spread and suppression (Dunn et al., 2017) in estimating WUI exposure risk.

Spatial metrics like building density, vegetation cover, and proximity to large areas of wildland fuels are used to identify buildings or neighborhoods that are vulnerable to wildfire (Gibbons et al., 2012, Bar-Massada et al., 2013, Radeloff et al., 2018). A growing body of research on the factors influencing home loss uses GIS focal analyses (within a neighborhood instead of at a single

pixel) to evaluate which landscape factors and at what scales are most predictive of home loss given exposure to wildfire (Price & Bradstock, 2013, Knapp et al., 2021, Syphard et al., 2021). Such models complement WUI classification methods with quantitative estimates of building loss risk, but they do not address the likelihood of WUI exposure to wildfire. Price et al., (2015) used binomial regression models to predict the probability of a fire spreading to the WUI by analyzing fire spread along artificial lines that connect random fire ignition points to potential WUI blocks and suggested that fire weather had the strongest influence, followed by the percentage of the artificial lines that were forested, the distance between a WUI and the approaching fire, and time since last fire.

Fire spread simulation models that account for weather, fuels, and topography as driving factors are also widely used by fire managers to inform suppression location, effort, and timing of evacuations. Two relevant examples in the USA are FARSITE (Finney, 2004, also called "near term fire behavior") and FSPro (Finney et al., 2011b), which are both integrated with the Wildland Fire Decision Support System (Calkin et al., 2011, Noonan-Wright et al., 2011). Haas et al., (2014) simulated thousands of random ignition locations spreading under extreme fire weather conditions on the Colorado Front range to identify potential ignition locations that have the highest potential to impact populated WUI areas. Ramirez et al., (2019) used stochastic fire simulation results to estimate the time and

Wei et al. Fire Ecology (2023) 19:30 Page 3 of 18

probability of a wildfire reaching WUI or other assets that trigger suppression decisions. While valuable, fire spread predictions depend on analyst expertise when setting parameter to calibrate observed fire behavior, and often provide limited ability to simulate the effects of fire suppression.

A separate but related body of research seeks to understand why fires cease to spread using empirical analysis of fire perimeters and characteristics of the landscape. One approach trains machine learning models of containment probability with random sample points from fire perimeters as successes and from fire interiors as failures (O'Connor et al., 2017). Another common approach is case-control logistic regression with successes sampled along the perimeter and paired failures extracted at points along a transect toward the fire interior (Narayanaraj and Wimberly, 2011, Macauley et al., 2022). Spatial predictions from these models may prove useful for rating the level of suppression opportunity between a wildfire and a WUI community.

In this study, we developed a novel focal analysis method to quantify how wind, landscape conditions, and fireline construction between a fire perimeter and nearby WUI blocks impacted WUI fire exposure probability. We designed spread funnels with a consistent shape to capture the potential fire spread toward WUI blocks. We used historical fire records to reflect the real-world fire and suppression situations in the western US. We trained random forest (RF) models to evaluate the importance of factors influencing WUI exposure probability. Our approach can be used to identify WUI communities with higher fire exposure probabilities during a fire incident, and help managers estimate the effects of suppression decisions.

Data and methods

Landscape conditions, wind, and suppression efforts may jointly influence WUI fire exposure probability during a wildfire incident. We leverage a set of historical fire data to study those influences.

Wildfire/WUI samples

We collected spatial data including fire perimeters, firelines, and nearby parcels of land classified as WUI for eleven large wildland fires (Fig. 1) in the western US from 2017 to 2018. These fires were selected, in part, based on the availability of data needed for our analysis and represent a wide range of environmental conditions. Fire perimeters from the GeoMAC historical fire

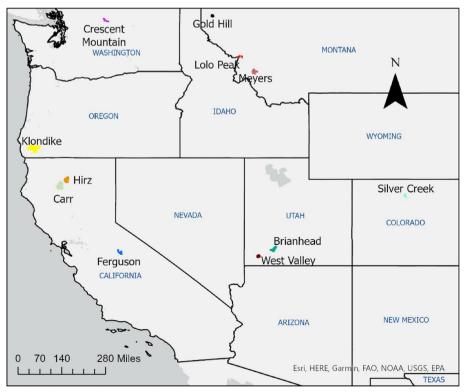


Fig. 1 The locations and final footprints of the eleven fires selected in the analysis

Wei et al. Fire Ecology (2023) 19:30 Page 4 of 18

Table 1 The list of eleven fires and the number of WUI blocks studied for each fire. The number of fire perimeters selected from each fire to pair with each WUI block varies depending on their locations

Year	Fire name	Fire size (ha)	Number of WUI blocks	Total area of WUI included in analysis (ha)
2017	Brianhead	29,013	17	1059
2017	Gold Hill	2641	1	1
2017	Lolo Peak	45,939	24	7490
2017	Meyers	51,987	4	76
2018	Carr	92,936	37	15,574
2018	Crescent Mountain	21,332	3	833
2018	Ferguson	39,186	22	51,670
2018	Hirz	18,800	1	75
2018	Klondike	84,011	36	38,085
2018	Silver Creek	8145	1	960
2018	West Valley	4769	1	4

database (Geospatial Multiagency Coordination Center, n.d) were used to capture fire progression throughout the duration of each fire incident. Fireline data were acquired from the National Interagency Fire Center (NIFC, https://data-nifc.opendata.arcgis.com/). We only included on-the-ground lines labeled as "completed" that were built with burnout operations, dozers, hand crews, plows, or that occurred along roads. The fireline data did not include other detailed line attributes such as line width or construction effort.

We downloaded a national WUI map derived from census blocks data (Radeloff et al., 2005). We dissolved the boundaries of adjacent WUI polygons to create continuous "WUI blocks." We selected blocks within

1600 m of the final fire perimeters for use in the analysis. Table 1 lists the number of blocks that were identified for each fire. A dependent variable "isImpact" (Table 2) is used to track whether the final fire perimeter intersected the selected WUI block (isImpact = 1) or not (isImpact = 0). Increasing the buffer distances of final fire perimeters would increase the number of WUI blocks not-impacted in our sample but would not change the number of impacted WUI blocks.

The progression of each fire is represented by multiple fire perimeters at daily or multi-day time steps (an example given in Fig. 2). In this study, we randomly included no more than twelve recorded fire perimeters from each fire. For fires with fewer than twelve

Table 2 List of variables selected for use within the random forest models. N/A indicates that there is no need to summarize the measurement across all funnels in each cluster as the measurements will be identical across funnels in the same cluster

Explanatory variables	Description of explanatory variables	Method to summarize for multiple funnels in a cluster
isImpact	Dependent variable denoting whether a fire would finally spread into a WUI block (isImpact=1) or not (isImpact=0)	N/A
Gust	The average hourly wind gust speed within a 2-day window before the earliest recorded day when a fire entered a WUI block; in case fire did not impact a WUI, a 2-day window before the final fire containment is used; measured by meter per second.	N/A
leakRatio_max	Max leak ratios calculated for all funnels in the same cluster.	Max
PCLMean	Average funnel-mean-PCL in the same cluster.	Average
rdDensityClosest	Road densities in the funnel of the closest fire perimeter to a WUI. It is calculated as the total number of raster cells within the 120-m buffers of all roads divided by the total number of raster cells in the funnel.	Closest
SDIClosest	Average SDI in the funnel formed from the closest fire perimeter to each WUI in each cluster.	Closest
WUIDist	Distance between the earliest recorded fire perimeter and a WUI block; measured by meters.	N/A
WUISize	The size of the WUI block of interest; measured by acres.	N/A

Wei et al. Fire Ecology (2023) 19:30 Page 5 of 18

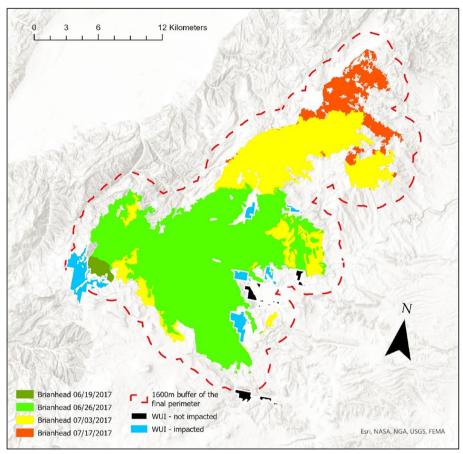


Fig. 2 Daily fire footprints for the Brianhead fire and its surrounding WUI blocks. Only those WUI blocks intersecting the 1600-m buffer of the final fire footprint are included in the analysis

recorded perimeters, all perimeters were included. Within included fire perimeters, we only chose those that were at least 1200 m from a WUI block. This process creates a "cluster" of fire perimeters (explained in Fig. 3 later) for each WUI block of interest.

Suppression difficulty and control potential

To provide metrics reflecting the relative effectiveness and difficulty to construct fireline between a fire perimeter and a WUI block, we used two raster map products from the decision support analytics used by the USDA Forest Service Risk Management Assistance (RMA) Program (Calkin et al., 2021) that were customized for use on the studied incidents. Suppression difficulty index (SDI) is a relative measure of how hard it is to suppress a fire with ground resources. It is calculated by dividing an "energy behavior index" by the sum of indices that influence fireline construction efficiency, such as accessibility, penetrability, and mobility (see Rodríguez y Silva

et al., 2020, 2014, O'Connor et al., 2016). The energy behavior index is calculated using flame length and heat per unit area (detailed equations presented in Rodríguez y Silva et al., 2020). Potential control location (PCL) suitability is the estimated probability of a fire ceasing growth based on the statistical association of binary control outcomes from historical fire perimeters with landscape characteristics relevant to fire suppression such as elevation, slope, aspect, fuel model, vegetation type, roads, accessibility, and suppression difficulty (O'Connor et al., 2017). PCL does not account for the presence or absence of firelines. Both SDI and PCL were originally calculated and represented as 30-m raster maps and are widely used to support containment strategy development during fire incidents (Calkin et al., 2021) as well as pre-fire planning (Dunn et al., 2020, Thompson et al., 2021). To improve the processing speed when using those maps, we resampled both layers into 120-m resolution.

Wei et al. Fire Ecology (2023) 19:30 Page 6 of 18

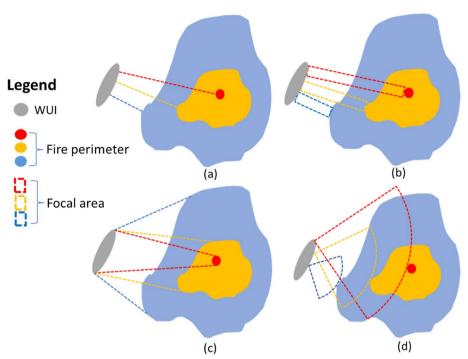


Fig. 3 Illustration of four potential focal area selections: **a** creating straight lines to connect a WUI block and corresponding fire perimeters; **b** forming rectangular focal areas; **c** connecting boundaries between fire perimeters and WUI as focal zones; **d** creating funnel-shaped focal analysis areas (used in this study). In **d**, the three funnels form a "cluster." Because the main axis of each funnel matches the unique shortest-distance-line between a WUI and a specific fire perimeter, the three funnels may overlap partially. The smallest funnel (delineated by dashed blue lines) in **d** is formed by the "closest" fire perimeter in the cluster

Focal analysis

Alternative focal area selections

A necessary step toward analyzing the WUI fire impact probability is to select the landscape (focal area) between a fire perimeter and WUI block for analysis. Different focal area designs may each have its strengths and weaknesses. We illustrate four options in Fig. 3. A simple approach would be connecting WUIs and fire perimeters with straight lines and analyzing the landscape conditions along each line segment (Fig. 3a). This design was implemented by Price et al., (2015) by connecting WUI and fire with random lines. However, studying spatial patterns only along those lines is narrowly focused. An improvement would be buffering those lines to create rectangular-shaped analysis zones to include broader potential fire spread areas (Fig. 3b). Another potential design is to connect the same-side boundaries between each fire perimeter and a WUI block (Fig. 3c). This design, however, would create focal zones with different shapes and pose challenges in comparing certain spatial metrics between zones. This design also tends to create relatively "narrow" focal areas from smaller fire footprints (Fig. 3c). In this study, we choose to build consistent "funnel" shaped (Fig. 3d) focal areas instead. The size and orientation of each funnel may vary, but all funnels would have an identical shape. The length of the "wider" edge of a funnel increases as the distance between a fire and WUI increases, which helps capture a broader area of possible fire spread areas toward a WUI when a fire is far away from it. Funnels pointing toward the same WUI from multiple perimeters of the same fire form a funnel cluster. We treat each funnel cluster as one sample. Synthesizing the landscape metrics from multiple funnels in the same cluster could better capture the overall landscape and fire suppression conditions along the potential fire spread paths toward a WUI.

Analyze funnel-shaped focal areas

As we discussed before, for each WUI/fire perimeter pair, we selected a funnel-shaped focal analysis area to study landscape conditions (PCL, SDI, roads, etc.) and fireline constructions within it. The two main steps for this approach are explained below and illustrated in Fig. 4.

- Step 1: Identify the shortest Euclidean distance between each pair of studied WUI/fire perimeter.
- Step 2: Create a fire impact funnel (or funnel) along the shortest distance identified from Step 1 (Fig. 4)

Wei et al. Fire Ecology (2023) 19:30 Page 7 of 18

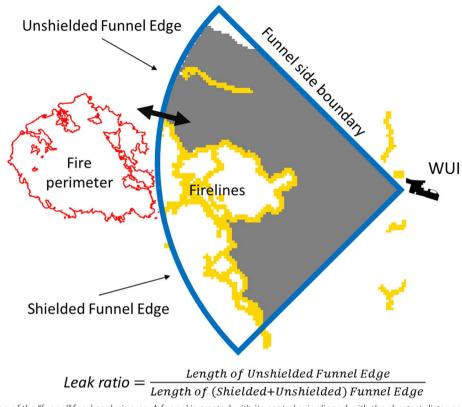


Fig. 4 An illustration of the "funnel" focal analysis area. A funnel is created with its central axis aligned with the shortest distance line connecting a WUI and a fire perimeter; each funnel has a fixed 90-degree opening angle. The funnel "leak ratio" is calculated as the length of the funnel edge not "shielded" by any completed firelines divided by the total length of the funnel edge along the wider side

using a raster-based algorithm (algorithm and examples provided in Appendix 1).

Each funnel is a quarter-circle-shaped focal area designed to delineate the landscape where a fire may likely spread across before impacting WUI. The funnel's axis will align with the shortest straight line connecting a fire perimeter to a WUI block. In each funnel, we analyzed how the firelines completed during suppression may "shield" a WUI from the threat of an upcoming fire. For this purpose, we calculated a measurement named "funnel leak ratio" or "leak ratio" as described in Appendix 1. In cases where firelines totally shield a WUI block from a nearby fire, the leak ratio of that funnel is zero; without any fireline, the leak ratio will be one; for a partially shielded funnel by firelines, the leak ratio could take any value between zero and one (e.g., ~ 0.33 in Fig. 4). The leak ratio was calculated based on completed firelines within a funnel instead of only firelines held successfully as we are interested in analyzing the potential impact of total fireline construction efforts (e.g., a breached fireline may still be able to slow fire spread). Besides the leak ratio, we also calculated other landscape metrics within each funnel, e.g., mean SDI, PCL, fireline density, and road density. These landscape measurements (except the leak ratio) were calculated in each funnel regardless of the fireline construction effort. Each funnel (e.g., in Fig. 4) provides a snapshot in time of the landscape between a WUI and an approaching fire perimeter. Multiple funnels are constructed for multiple perimeters from the same fire to form a cluster for the same WUI to provide multiple snapshots of landscapes between a WUI and an approaching fire (Fig. 3d).

Gust impact

The impact of wind to fire and suppression is complicated and highly variable (Keeley & Syphard, 2019). For each of the 11 fires studied, we collected wind records from the closest fifteen weather stations (Appendix 2, Table 3) within a 100-km radius of each fire from the Synoptic Data portal (https://developers.synopticdata.com/mesonet/v2/). To represent wind occurring during active fire spread periods, we used wind records

Wei et al. Fire Ecology (2023) 19:30 Page 8 of 18

from ten o'clock local time in the morning until nine o'clock at night. It would be ideal if we could have collected more precise wind records matching the time when fire encountered fireline, but this type of data was not available. We chose to use the average hourly gust speed across the 2-day period prior to the day that a fire impacted a WUI block, or before the fire was contained (for not-impacted WUI blocks), as a coarse-scale explanatory variable to study WUI fire exposure probability. Gusts also have strong positive correlations with sustained wind speed (Fovell & Gallagher, 2018). The experiments from aggregating gust speeds for alternative durations are presented in Appendix 3 and Figure 9.

Explanatory variables

After testing many candidate explanatory variables, we selected seven of them to build RF models (Table 2). Some of the variables have been discussed earlier in this paper. For example, we calculate the mean PCL, SDI, and road density in each impact funnel. The variables not yet mentioned include the distance between the earliest recorded perimeter of each fire and each studied WUI block (denoted "WUIDist") and the size of each WUI block (denoted "WUISize"). We also selected the maximum leak ratio across all funnels in the same cluster (denoted leakRatio_max in Table 2). When screening explanatory variables, we reference the correlation heatmaps between variables (Figure 10 in Appendix 4) by selecting variables generally with higher correlation with the predicted variable (e.g., isImpact) but lower correlations with the other explanatory variables.

Building and testing random forest models

Random forests use a combination of decision tree predictors with each tree built on a set of random samples (Breiman, 2001). We treat each funnel cluster as a sample used to train the RFs; each cluster may include one or multiple impact funnels associated with the same WUI block and fire.

We use the *python 3.0 sklearn.RandomForestClassifier* module to train and test RF models based on the seven selected explanatory variables. Hyperparameters were used to control the RF learning process. A cross-validation method from the *sklearn's RandomizedSearchCV* function was implemented to test different combinations of hyperparameters. Based upon those test results, we chose to use 300 tree-classifier results to construct each RF; the minimum number of samples in each leaf is set to one, and the maximum depth of each tree is set to six. All explanatory variables were normalized prior to being used to create the RF models; thus, their values ranged between zero and one.

To better understand the RF models' performance, we adopt a Monte Carlo random sampling process to build 200 RF model replications; in each replication, we randomly split all clusters into a training and a testing data set, with each training set comprised of 80% of the total clusters and the remaining 20% of clusters designated as the corresponding test set. This method builds one RF tree using the training set and tests that RF against the remaining test data set. Training and testing multiple RFs allow us to estimate both the mean and variation of the RF models when they were used to study the importance of each expandatory variable on each fire's impact on WUI blocks.

Results

Below we provide test results describing the input data to the RF models and metrics reflecting the RF models' performance.

Data description

We create histograms of the seven variables (Table 2) associated with both the "impacted" and "not-impacted" WUI blocks (Fig. 5). Observations show that impacted WUI blocks are generally associated with higher maximum leak ratios; in contrast, a maximum leak ratio of zero for all funnels within a cluster (the WUI block is shielded by firelines in all the funnels) is often associated with not-impacted WUI blocks. PCL is a measure of the probability that a fire will cease at any location irrespective of suppression effort. An average PCL > 0.4 in a funnel cluster is often associated with not-impacted WUI blocks. Both SDI and roads directly reflect fire suppression difficulty and opportunity, which are also important measurements in the "closest" funnel to the WUI block within each cluster. For the closest funnel, if its average SDI was less than 50 (about 30 percentile) or its road density was higher than 0.4, we more likely find the corresponding WUI blocks not impacted by a fire. A 2-day average of hourly wind gust faster than 4.5 m/s before a fire impacted a WUI appears to create a greater threat to the studied WUI blocks.

RF classification results

We evaluated the performance of RF models through cross-validation using different testing data sets to predict whether each WUI block would be impacted versus not-impacted by a fire (Fig. 6a). The average RF classification accuracy from cross-validation is about 86%. The "area under the ROC curve" (AUC) was calculated for each RF model; AUC measures the two-dimensional area underneath the receiver operating characteristic (ROC) curve from (0,0) to (1,1) and represents model performances across all classification thresholds. The mean

Wei et al. Fire Ecology (2023) 19:30 Page 9 of 18

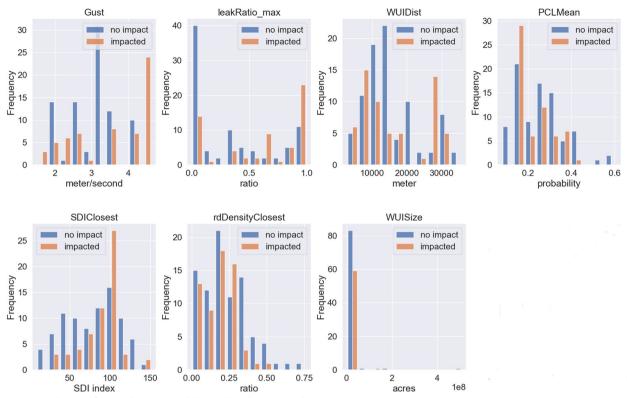


Fig. 5 Distribution of the explanatory variables in either the "impacted" or "not-impact" groups

AUC value of the RF models is 86%. To test each of the RF classifiers individually using different random test datasets, a two-by-two confusion matrix was constructed representing the dispositions of the RF classifications (Fawcett, 2006). The confusion matrix summary generated from 200 RFs (Fig. 6b) had an average true-negative rate (TNR) rate of ~0.88 and average true-positive rate (TPR) rate of ~0.83, suggesting the RF models slightly under-predict the positive (impacted) condition from the testing data sets used for cross-validation (Fig. 6b).

From the predictors' relative importance graph (Fig. 6c), we found that three explanatory variables show relative importance levels above 0.10. Of those variables, gust is ranked as the most important variable, with an average relative importance of 0.34. The max leak ratio in each funnel cluster is the second most important explanatory variable, with an average relative importance level of 0.16. The distance between the earliest fire perimeter in the cluster and the WUI block has an average relative importance of 0.15. The importance of the remaining four variables is close to or slightly below 0.1. For example, the cluster average of PCL has a relative importance level of 0.1, similar to the road density and SDI measurements in the closest funnel of each cluster. The size of WUI block shows a smaller average relative importance of 0.07.

We calculated and displayed the average partial dependence (PD) measurements from the 200 tested RF models (Friedman, 2001; Molnar 2022) for each of the seven explanatory variables (Fig. 7). PDs reported by *sklearn* across a range of values of each explanatory variable provide the average prediction of WUI fire exposure probability if all data points are set to assume the value of the analyzed variable. Although PD values fluctuate for some explanatory variables due to randomness, noise, and smaller sample sizes within certain data ranges, the general trends of the PD curves could still reflect the impact of those explanatory variables on WUI fire exposure probabilities.

PD analyses show that gust values above the eightieth percentile are associated with dramatically increased WUI fire exposure probability. Increasing values of leakRatio_max are generally associated with an increase in WUI fire exposure probability, especially when the leakRatio_max is above 0.6. The ability to plan and implement suppression actions if a fire can be discovered when it is still far away from a WUI, represented by larger values of WUIDist, is associated with lower WUI fire exposure probability. Increasing road density is also correlated with lower WUI fire exposure probability. As expected, higher SDI and lower PCL are associated with higher WUI exposure probability. Note that both SDI and PCL are composite measurements

Wei et al. Fire Ecology (2023) 19:30 Page 10 of 18

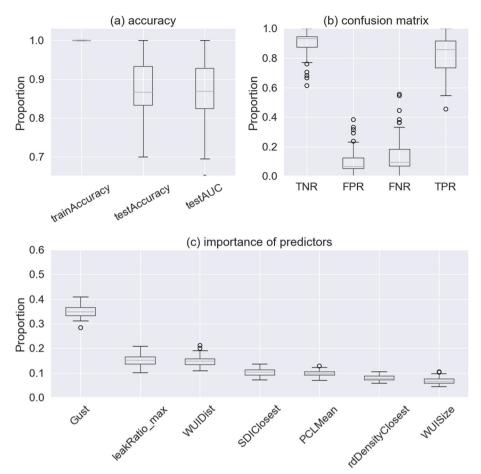


Fig. 6 RF models performances through random trainings and testing data sets. **a** Classification accuracy. **b** The confusion matrix (TNR, true-negative rate; TPR, true-positive rate; FPR, false-positive rate; FNR, false-negative rate). **c** The relative importance of predictors

reflecting vegetation and fuel conditions, accessibility, and fire cessation probability (PCL), or suppression challenges (SDI). Increasing road density in the closest funnel also decreased the probability of a WUI being impacted by a fire. In addition, larger WUI blocks may be more prone to fire impact. The overall trends shown in the PD graphs are consistent with our initial expectations.

Appendix 5 shows another RF model retaining only the three most important explanatory variables. Although removing four of the seven explanatory variables did not meaningfully decrease the prediction accuracy, we found higher levels of fluctuation of the PD curves, which make it more challenging to understand how each individual explanatory variable would influence the WUI fire exposure probability.

Discussion and conclusions

In this study, we constructed clusters of funnel-shaped focal analysis units between WUI polygons and active fires and analyzed factors that may influence WUI fire exposure probabilities. Centering our analyses on fixed-shape funnels allowed us to calculate and compare landscape characteristics among fires and at-risk WUI blocks. We quantified the potential "shielding" effect from building firelines in each funnel using a calculated leak ratio. We assume fireline construction effort will help mitigate WUI fire exposure risk either by containing or slowing fire spread. We built and tested a set of RF models based on randomly splitting training and testing datasets. We selected seven explanatory variables and examined their impact using predictor importance and PD plots. Test results show that using RF models with those predictor variables achieved reasonable classification accuracy for WUI fire impact outcomes. The RF models could be used to compare fire exposure probabilities between different WUI blocks under different wind and suppression scenarios to help managers identify the most vulnerable WUI blocks during a fire event. Our analyses show that funnel-specific landscape conditions such as road

Wei et al. Fire Ecology (2023) 19:30 Page 11 of 18

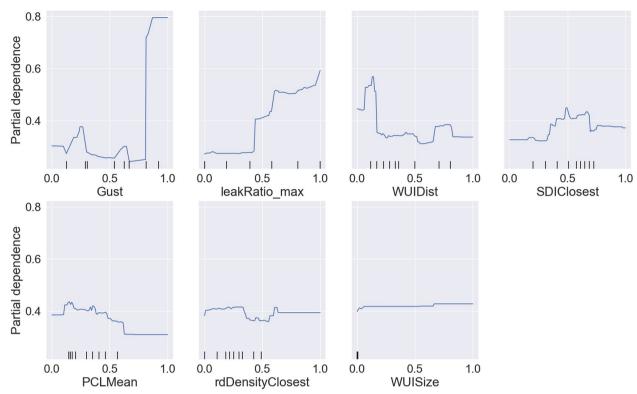


Fig. 7 Partial dependence (PD) plot for the seven explanatory variables, each normalized between zero and one and represented in the *x*-axis, versus the estimated probability of a fire impacting a WUI block, represented in the *y*-axis. Each point on the PD plot is the probability in favor of the "impacted" class across all observations. The nine ticks along the *x*-axis split the samples into ten groups with equal number of samples in each group (or deciles). Some ticks are clustered together and represent denser samples within the range

density, SDI, PCL, fireline amount and arrangement, distance between the initial fire discovery location and WUI block, and WUI block size are likely to impact the WUI fire exposure probability during a fire event.

This study accounted for the impact of wind gusts in the RF models using the average hourly wind gust speeds within a 2-day window prior to a fire's impact on a WUI block (for impacted WUI blocks), or prior to a fire being contained (for not-impacted WUI blocks). The RF model results indicated that higher gust speeds (positively correlated with higher sustained wind speeds) were the most important factor in increasing WUI fire exposure probability. Higher gust speeds may result in more severe fire behavior as well as associated firefighter safety concerns that can limit effective fireline construction and use. Due to spatial and temporal limitations in data coverage, it is often difficult to reconstruct wind impacts on fire spread and suppression along specific firelines, fuel breaks, or natural barriers (roads, ridge lines, etc.) Terrain may also greatly impact wind directions and speeds, which may have both positive and negative impacts on fire suppression. Future research may benefit from designing and testing additional wind measurements with better data.

Besides wind and firelines (reflected by the calculated leak ratio), the distances between the initial fire location when it was first recorded and the WUI blocks at risk were critical in determining WUI fire exposure probability. This is likely for two reasons: (1) fire is inherently less likely to spread a longer distance than a shorter distance and (2) a longer distance likely allows additional time and opportunity for managers to plan and implement suppression actions. In addition to the distances between initially recorded fire perimeters and WUI blocks, both the road density and SDI in the "closest" funnel to a WUI block show moderate impacts on WUI fire exposure probability. Roads likely play at least two functions—serving as pre-constructed fire breaks as well as facilitating access to the fire for personnel and equipment. The SDI metric represents the averaged suppression difficulty levels as a fire approaches a WUI block; thus, we can expect that lower SDI is associated with lower WUI fire exposure probability.

Wei et al. Fire Ecology (2023) 19:30 Page 12 of 18

The fire impact funnels provide an intuitive and consistent spatial delineation that allowed us to calculate landscape spatial metrics along likely fire spread pathways before a fire reaches a WUI block. We believe it represents a logical and intuitive design that is easy to implement, but we also acknowledge it is not necessarily a perfect spatial delineation of focal analysis areas. For example, setting up the opening angle of a funnel is somewhat arbitrary. We choose a 90-degree angle because it is easy to program on a raster map. We could either increase or decrease the funnel angle to account for broader or narrower areas across which a fire may spread toward a WUI block. There are also other shape options for the focal analysis zones, e.g., line, circle, ellipse, rectangle, or a combination of multiple shapes. Designing, testing, and comparing different focal zone designs, and potentially different methods to calculate landscape metrics, could improve model performance.

The impact of landscape conditions on WUI fire exposure is complex and thus will continue to provide interesting fodder for future research, particularly as the data to support such research improves. Beyond incorporating additional spatial metrics as explanatory variables to build RF models, another method that might be used to examine WUI fire exposure probabilities would be to use images to directly train convolutional neural networks (CNN). Images could be clipped out into consistent shapes such as squares, or by fire impact funnels, etc. CNN models can be trained to recognize more complex spatial patterns of fireline amount, orientations, shapes and locations, SDI or PCL, and roads, and learn how those patterns may jointly influence WUI fire exposure probabilities. Compared to the RF models tested here, CNN models can use images as direct inputs instead of relying on precalculated landscape spatial measurements such as leak ratio or average PCL, etc. Building CNN models, however, may require substantial effort to create labeled images and innovative model designs to utilize the information carried by those images.

Collecting well-synchronized spatial and temporal fire data is time-consuming and challenging. This study built and tested RF models using the focal analysis concept based on moderate sample sizes. Test results identify reasonable relationships between the WUI fire exposure probability and seven explanatory variables. Collecting more fire samples and adding more data layers, such as aerial retardant or water drops, fuel break locations, previous fires, or fireline types (dozer, handline, burnout, etc.), could also improve our ability to study factors influencing WUI fire exposure probabilities.

Appendix 1

Construct fire impact funnels and calculate funnel leak ratio

The axis of each fire impact funnel orientates along either the South, North, East, West, Northeast, Northwest, Southeast, or Southwest direction (denoted as S, N, E, W, NE, NW, SE, SW). For example, if a funnel axis has an azimuth direction of 185-degree, we will classify it as "North" facing or "N". Starting from the WUI end, we will add raster cells one by one into the impact funnel following a predefined sequence (Fig. 8a). During the process of building a funnel, we iteratively call equations (Eq.1) and (Eq.2) to add raster cells (indexed as c) into the funnel by assigning either 0 or 1 value to a variable x_c to track whether direct fire spread path without encountering any firelines can be found from cell c to the targeted WUI. If this path exists, we set $x_c = 1$; otherwise, $x_c = 0$. Cells not part of the funnel will have $x_c = null$. Since this algorithm is designed to use raster maps as inputs, all the input map layers such as firelines, WUI, roads, SDI, and PCL need to be converted to raster maps first.

$$x^{\text{WUI}} = 1 \tag{1}$$

$$x_c = \begin{cases} 1, & \text{if } \sum_{c'} x_{c'} \times (1 - \text{line}_{c'}) \ge 1 \ \forall c \text{ added to the funnel} \\ 0, & \text{otherwise} \end{cases}$$
 (2)

- Set $x_c = \text{null for all cells}$ as the initial state.
- Variable x^{WUI} is set to 1 for the raster cell(s) representing the possible fire impact point(s) to the corresponding WUI.
- An impact funnel will be formed by keeping adding new cells c into the funnel. c' denotes cells directly adjacent to cell c from the WUI side.
- Parameter line_{c'} denotes whether firelines exist in cell c'. Zero denotes no fireline in cell c'; 1 denotes fireline existing in c'.

For example, assuming we have a raster map with row index (denoted as row_ID) and column index (denoted as col_ID) for each cell; $(row_ID=0, col_ID=0)$ represents the location of the top-left conner of the raster map; row_ID increases by moving downward and col_ID increases by moving rightward across the map. If we assume a fire was discovered from the north of a studied WUI (Fig. 8), for a cell at the location (row_ID, col_ID) , at most three directly adjacent cells c' connected to it from the WUI side of cell c with their locations indexed as $(row_ID + 1, col_ID - 1)$, $(row_ID + 1, col_ID)$ and $(row_ID + 1, col_ID + 1)$. The sequence of assigning values to cell c shown in Fig. 8a

Wei et al. Fire Ecology (2023) 19:30 Page 13 of 18

guarantees that variable x_c is assigned a value of either zero or one (Eq. 2) only after $x_{c'}$ for all its three directly adjacent cell c' from the WUI side of the funnel have already been calculated. $x_{c'}$ is assigned a value of zero under two scenarios according to Eqs. 1 and 2: (1) $x_{c'}$ of all the three adjacent cells c' have been assigned a value of zero, or (2) fireline exists in all the three adjacent cell c'.

Using the example in Fig. 8, after a funnel is built, we can tally the value of x_c for cells at the wider end of the funnel.

Any cell c along the wider side of the funnel edge with $x_c = 1$ is identified as a cell that can spread fire to the studied WUI following a combination of vertical and diagonal spread directions (Fig. 8). The number of cells along the edge with $x_c = 1$ divided by the number of cells along the wider end of the *funnel* edge is called "leak ratio" (Table 2). If firelines can separate a WUI and a fire perimeter in a *funnel*, the *leak ratio* of the *funnel* would be zero; in other cases, the leak ratio would be between zero and one.

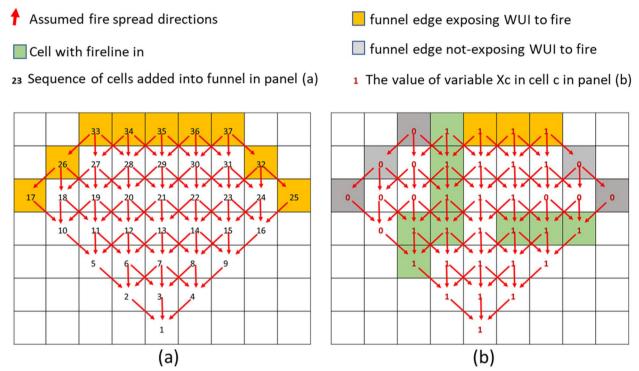


Fig. 8 In this demo, we assume fire would spread toward a WUI from the north end of the *funnel* and a WUI is located at the south end of the *funnel*. **a** The sequence of cells being added into the *funnel*. Panel **a** also shows a case that fire from the wider end of the funnel edge can spread to the WUI without encountering any fireline. **b** A case with firelines built in the funnel. The "leak ratio" after considering the effect of fireline is about $4 \div 9 = \sim 0.44$ in Panel b.

Wei et al. Fire Ecology (2023) 19:30 Page 14 of 18

Appendix 2

Table 3 The weather station codes used for each fire through the *Synoptic Data portal*.

Fire name	Weather station IDs	
Brianhead	BHCU1; BRHU1; COOPPAGU1; CTMU1; CVYU1; PGLU1; PNGU1; UCC19; UCC20; UTR20; YKEU1;	
Carr	BDYC1; CLCC1; CQ098; CTBSN; HDLC1; LSHC1; MMOC1; OBRC1; RFSC1; SFKC1; SHDC1; SLFC1; TCAC1; TRWC1;	
Crescent Mountain	COOPMZAW1; COOPSTEW1; MAZ22; RAIW1; RDHW1; STRW1; USGSHY657; USGSHY666; WAP55; WAP67;	
Ferguson	C1522; CNFC1; CQ160; CQ168; D0826; EPWC1; GFTC1; JSDC1; MAGC1; MBBC1; MIAC1; WCFC1; WWNC1; YWC1; YWAC1;	
Gold Hill	BANM8; BIFM8; COOPLIBM8; COOPTRSM8; KS59; LBBM8; MTYAK; S599; TROM8; TS259; TT166; TT481; YKAM8; ZONM8;	
Hirz	BNKC1; CTANT; CTVOL; GISC1; GRDC1; HRZC1; LSHC1; SLFC1; SMSC1; UP595; UP627; UP670;	
Klondike	AGFO3; E4806; E7191; K3S8; ODT23;	
Lolo Peak	COOPSTEM8; E0591; ITD28; LOLM8; LPSI1; MOMM8; MTM01; SMTM8; STVM8; TS934;	
Meyers	BRLM8; CLVM8; COOPSLAM8; DALM8; DEEM8; GPRM8; ITD87; MTGTL; PHGM8; PTNM8; SKAM8; SUAM8; TEPM8; TT282;	
Silver Creek	ARPC2; BUFC2; C9561; COLC2; COOPKMLC2; CSU61; GSPC2; KC07; LYNC2; PCPC2; RESC2; YCAC2;	
West Valley	CAVU1; CDCU1; COOPNHRU1; FG016; GAPU1; LGFU1; LPRU1; TS716; UTBLK; UTTP3; VEYU1; WRRU1;	

Appendix 3

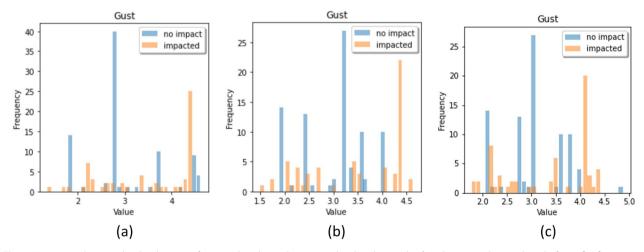


Fig. 9 Histogram showing the distributions of average hourly wind gust speed within the **a** 1-day, **b** 2-day, or **c** 3-day windows before a fire first impacted a WUI block (for impacted WUI) or before a fire was finally contained (for not-impacted WUI). Results suggest that using a 2-day window provides the cleanest classification results based on the given samples

Wei et al. Fire Ecology (2023) 19:30 Page 15 of 18

Appendix 4Additional descriptive statistics of the explanatory variables

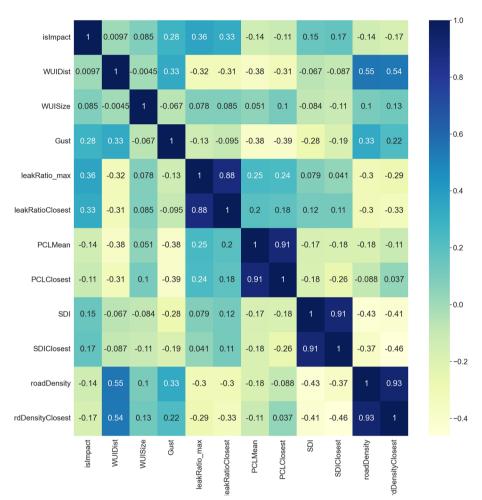


Fig. 10 Correlation heat map between candidate explanatory variables

Wei et al. Fire Ecology (2023) 19:30 Page 16 of 18

Appendix 5

Figure 11 shows the performance from another set of RF models using only the three most important explanatory variables (3-variable models). Comparing with the 7-variable RF models, the 3-variable models provide just similar prediction accuracy (Figure 11a, b). As expected, the

relative importances of the three explanatory variables are higher than the same variables from the 7-variable model given in the main text (Figure 11c). The PD graphs from the 3-variable model follow the same pattern as the same variables in the 7-variable models, but with higher levels of fluctuation (more jagged, or less smooth).

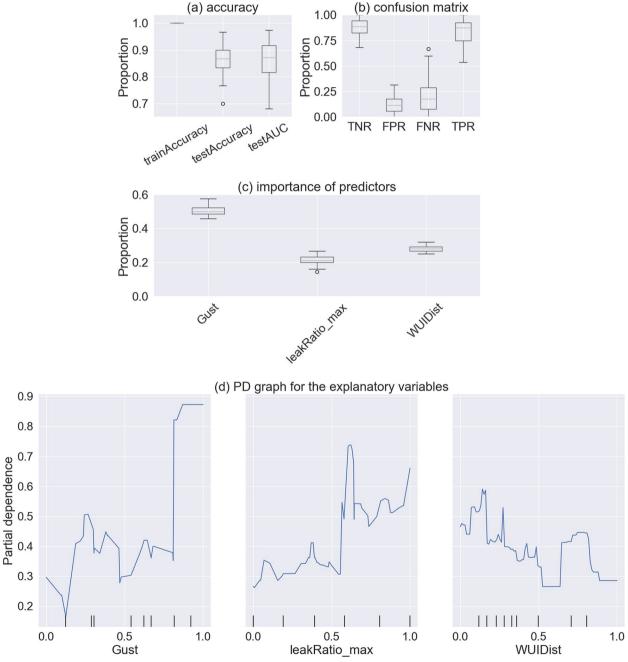


Fig. 11 RF model performance through random training and testing data sets for the 3-variable model. a Classification accuracy. b The confusion matrix. c The importance of explanatory variables. d The PD graphs for the three explanatory variables

Wei et al. Fire Ecology (2023) 19:30 Page 17 of 18

Abbreviations

AUC Area under the ROC Curve
CNN Convolutional neural networks
NIFC National Interagency Fire Center
PCL Potential control location
PD Partial dependence
RF Random forest
RMA Risk Management Assistance

RMA Risk Management Assistance
ROC Receiver operating characteristic
SDI Suppression difficulty index
WUI Wildland urban interface

Acknowledgements

Chris Dunn from Oregon State University provided some recent GIS data used in the analysis.

Disclaimer

The findings and conclusions in this report are those of the author(s) and should not be construed to represent any official USDA or US Government determination or policy. This research was supported, in part, by the US Department of Agriculture, Forest Service. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US government.

Authors' contributions

YW, MT, and DC identified the research problem and obtained funding. YW and MT designed the funnel analysis method. YW wrote codes to conduct the analysis. YW drafted the manuscript. BG, JY, and EB reviewed the methods in data collection, data analysis, and random forest model formulations. CO provided key spatial data and suggested how to use them. All contributed editorial input during manuscript preparation and revision. The authors read and approved the final manuscript.

Funding

This project was funded by a joint venture agreement (18-JV-11221636-170) between the Colorado State University and the USDA Forest Service Rocky Mountain Research Station.

Availability of data and materials

The datasets used and/or analyzed here are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

The authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 6 June 2022 Accepted: 18 April 2023 Published online: 10 May 2023

References

- Abatzoglou, J.T., and A.P. Williams. 2016. Impact of anthropogenic climate change on wildfire across western US forests. *Proc Natl Acad Sci USA* 113 (42): 11770–11775.
- Bar, Massada A., A.D. Syphard, S.I. Stewart, and V.C. Radeloff. 2013. Wildfire ignition distribution modeling: a comparative study in the Huron-Manistee National Forest, Michigan, USA. *International Journal of Wildland Fire* 22: 174–183.
- Breiman, L. 2001. Random forests. Machine learning 45 (1): 5-32.
- Caggiano, M.D., T.J. Hawbaker, B.M. Gannon, and C.M. Hoffman. 2020. Building loss in WUI disasters: Evaluating the core components of the wildland–urban interface definition. *Fire* 3 (4): 73. https://doi.org/10.3390/fire3040073.

- Calkin, D.E., M.P. Thompson, M.A. Finney, and K.D. Hyde. 2011. A real-time risk assessment tool supporting wildland fire decision making. *Journal of Forestry* 109 (5): 274–280.
- Calkin, D.E., C.D. O'Connor, M.P. Thompson, and R.D. Stratton. 2021. Strategic wildfire response decision support and the risk management assistance program. *Forests* 12 (10): 1407. https://doi.org/10.3390/f12101407.
- Dunn, C.J., M.P. Thompson, and D.E. Calkin. 2017. A framework for developing safe and effective large-fire response in a new fire management paradigm. *Forest Ecology and Management* 404: 184–196. https://doi.org/10.1016/j.foreco.2017.08.039.
- Dunn, C.J., C.D. O'Connor, J. Abrams, M.P. Thompson, D.E. Calkin, J.D. Johnston, R. Stratton, and J. Gilbertson-Day. 2020. Wildfire risk science facilitates adaptation of fire-prone social-ecological systems to the new fire reality. *Environmental Research Letters* 15: 025001.
- Fawcett, T. 2006. An Introduction to ROC Analysis. *Pattern Recognition Letters* 27 (8): 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.
- Finney, M.A. 2004. FARSITE: Fire Area Simulator–model development and evaluation. *Research Paper* RMRS-RP-4 Revised. Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Finney, M.A., C.W. McHugh, I.C. Grenfell, K.L. Riley, and K.C. Short. 2011. A simulation of probabilistic wildfire risk components for the continental United States. *Stoch Env Res Risk A* 25: 973–1000.
- Finney, M.A., I.C. Grenfell, C.W. McHugh, R.C. Seli, D. Trethewey, R.D. Stratton, and S. Brittain. 2011. A method for ensemble wildland fire simulation. *Environ Model Assess* 16: 153–167.
- Fovell, R..G., and A. Gallagher. 2018. Winds and gusts during the Thomas fire. *Fire* 1: 47. https://doi.org/10.390/fire1030047.
- Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 1189-1232
- Geospatial Multiagency Coordination Center. n.d. Historic GeoMAC Perimeters; National Interagency Fire Center: Boise, ID, USA. Available online: https://data-nifc.opendata.arcgis.com
- Gibbons, P., L. van Bommel, A.M. Gill, G.J. Cary, and D.A. Driscoll. 2012. Land management practices associated with house loss in wildfires. *PLoS One* 7 (1): e29212.
- Haas, J.R., D.E. Calkin, M.P. Thompson, and M.A. Finney. 2014. Wildfire risk transmission in the Colorado Front Range, USA. *Risk Analysis* 35 (2): 226–240.
- Iglesias, V., Balch, J.K. and Travis, W.R. 2022. U.S. fires became larger, more frequent, and more widespread in the 2000s, *Science Advances*. https://doi.org/10.1126/sciadv.abc0020
- Keeley, J., and A. Syphard. 2019. Twenty-first century California, USA, wildfires: fuel-dominated vs wind dominated fires. Fire Ecology 15: 24. https://doi. org/10.1186/s42408-019-0041-0.
- Knapp, E.E., Y.S. Valachovic, S.L. Quarles, and N.G. Johnson. 2021. Housing arrangement and vegetation factors associated with single-family home survival in the 2018 Camp Fire California. *Fire Ecology* 17: 25. https://doi. org/10.1186/s42408-021-00117-0.
- Macauley, K., N. McLoughlin, and J.L. Beverly. 2022. Modelling fire perimeter formation in the Canadian Rocky Mountains. *Forest Ecology and Management* 506: 119958. https://doi.org/10.1016/j.foreco.2021.119958.
- Molnar, C. 2022. Interpretable Machine Learning: A guide for making black box models explainablehttps://christophm.github.io/interpretable-ml-book/.
- Narayanaraj, G., and M.C. Wimberly. 2011. Influences of forest roads on the spatial pattern of wildfire boundaries. *Int J Wildl Fire* 20: 792–803. https://doi.org/10.1071/WF10032.
- Noonan-Wright, E., Opperman, T.S., Finney, M.A., Zimmerman, G.T., Seli, R.C., Elenz, L.M., Calkin, D.E., Fiedler, J.R. 2011. Developing the U.S. wildland fire decision support system. *Journal of Combustion*. Article ID 168473. 14p.
- O'Connor, C.D., M.P. Thompson, and Rodríguez y Silva, F. 2016. Getting ahead of the wildfire problem: Quantifying and mapping management challenges and opportunities. *Geosciences*. 6: 35.
- O'Connor, C.D., D.E. Calkin, and M.P. Thompson. 2017. An empirical machine learning method for predicting potential fire control locations for pre-fire planning and operational fire management. *International Journal of Wildland Fire* 26: 587–597.
- Platt, R.V., T. Schoennagel, T.T. Veblen, and R.L. Sherriff. 2011. Modeling wildfire potential in residential parcels: a case study of the north-central Colorado Front Range. *Landscape and Urban Planning* 102 (2): 117–126.
- Plucinski, M.P. 2019. Fighting flames and forging firelines: Wildfire suppression effectiveness at the fire edge. *Current Forestry Reports* 5: 1–19.

Wei et al. Fire Ecology (2023) 19:30 Page 18 of 18

- Price, O., R. Borah, R. Bradstock, and T. Penman. 2015. An empirical wildfire risk analysis: The probability of a fire spreading to the urban interface in Sydney, Australia. *International Journal of Wildland Fire* 24: 597–606.
- Price, O., and R. Bradstock. 2013. Landscape scale influences of forest area and housing density on house loss in the 2009 Victorian bushfires. *PLoS ONE* 8 (8): e73421. https://doi.org/10.1371/journal.pone.0073421.
- Radeloff, V.C., D.P. Helmers, H.A. Kramer, M.H. Mockrin, P.M. Alexandre, A. Bar-Massada, V. Butsic, T.J. Hawbaker, S. Martinuzzi, A.D. Syphard, and S.I. Stewart. 2018. Rapid growth of the US wildland-urban interface raises wildfire risk. *National Academy of Sciences* 115 (13): 3314–3319.
- Radeloff, V..C., R..B. Hammer, S..I. Stewart, J..S. Fried, S..S. Holcomb, and J..F. McKeefry. 2005. The Wildland-Urban Interface in the United States. *Ecological Applications* 15 (3): 799–805 http://www.jstor.org/stable/4543395.
- Ramirez, J., S. Monedero, C.A. Silva, and A. Cardil. 2019. Stochastic decision trigger modelling to assess the probability of wildland fire impact. *Science of the Total Environment* 694: 133505.
- Rodríguez y Silva, F.R., J.R.M. Martínez, and A. González-Cabán. 2014. A methodology for determining operational priorities for prevention and suppression of wildland fires. *International Journal of Wildland Fire* 23: 544–554
- Rodríguez y Silva, F.R., C..D. O'Connor, M..P. Thompson, J.R..M. Martínez, and D.E. Calkin, 2020. Modelling suppression difficulty: Current and future applications. *International Journal of Wildland Fire* 29: 739–751.
- Stratton, R.D. 2020. The path to strategic wildland fire management planning. *Wildfire* 29 (1): 24–31.
- Syphard, A.D., H. Rustigian-Romsos, and J.E. Keeley. 2021. Multiple-scale relationships between vegetation, the wildland-urban interface, and structure loss to wildfire in California. *Fire* 4: 12.
- Thompson, M.P., B.M. Gannon, and M.D. Caggiano. 2021. Forest roads and operational wildfire response planning. *Forests* 12: 110. https://doi.org/10. 3390/f12020110.
- US Department of Interior and US Department of Agriculture. 2014. National Cohesive Wildland Fire Management Strategy. Available at https://www.forestsandrangelands.gov/strategy/ [Verified 31 May 2022]

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ▶ Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com